

Abstract of the Disclosure

0000070.061101
101190.000000

A document collection apparatus collects a document for a community from a network. Before
5 starting the collection, an initial document group is given as a starting point of the collection. A reference extraction unit extracts a reference from the initial document group. A next prospect determination unit determines a prospect which is a
10 prospective document to be collected next. A document collection unit collects the prospect to be collected next, and adds it to the collected documents. The document collection unit first collects documents evenly from inside the community
15 in the network, and then the next prospect determination unit determines the prospect to be collected next from inside and outside the community.